



Commercial Participation in OpenStack: Two Sides of a Coin

Yuxia Zhang, Beijing Institute of Technology

Hao He and Minghui Zhou, Peking University

This article provides a landscape of commercial participation in OpenStack, a large-scale open source software (OSS) ecosystem. We discuss how to achieve a balance between maximizing business profit and ensuring the long-term sustainability of OSS ecosystems.

Open source software (OSS) has dramatically influenced computing and society,¹ becoming the infrastructure of our daily life. As reported, more than 90% of software products contain OSS components.² The success of OSS 1.0 relies on the volunteers distributed all over the world, who bring to it open innovation, low cost, and high quality.

Sensing the business opportunities of OSS, numerous companies have built business models around OSS

ecosystems to achieve innovations,³ reduce costs,⁴ or generate revenue on complementary services.⁵ To achieve their goals in an OSS ecosystem, companies hire developers to contribute to the projects within that ecosystem.⁶⁻⁸ Many well-known OSS ecosystems such as the Linux kernel, Android, and OpenStack are developed mainly through collaborations of different companies. For example, companies contribute more than 90% of code on average in

each version of OpenStack.⁹ Brian Fitzgerald calls the OSS with extensive commercial participation OSS 2.0.¹⁰

Unlike self-motivated contributors, companies are always driven by business profits,¹¹ even when they are making contributions to OSS.^{6,9} Companies involved in OSS may try to control the development direction of the OSS projects to maximize business profits or simply withdraw all their employees from OSS to avoid investment failures. Therefore, while boosting software development, commercial participation also brings challenges and risks to the long-term development of OSS and the sustainability of critical open source ecosystems.



We conducted a series of studies^{7,9,12} to understand how companies participate with other companies, focusing on a large-scale OSS ecosystem, OpenStack. Following these studies, we used the code commit history of OpenStack to analyze how different development tasks are completed in OpenStack and conducted a survey to explore developers' perspectives on how company domination affects the OSS projects. In this article, by combining the findings from our prior work with the results on the task variation of companies and developers' feedback, we provide a landscape of commercial participation in OpenStack and propose lessons for the sustainability of OSS ecosystems. We hope this article can shed some light on shaping and sustaining OSS ecosystems as

FROM THE EDITOR

Welcome back to the "Open Source Expanded" column! We continue our theme on community open source. This article in the series by Minghui Zhou's research group at Peking University shows that communal ownership of software does not conflict with commercial interests. Rather, it is complementary and can work well when taken care of using good governance. Enjoy, and as always, stay safe and healthy and keep on hacking. —Dirk Riehle

commercial participation continues to increase in the future.¹³

THE PREVALENT COMMERCIAL PARTICIPATION AND ITS CHARACTERISTICS

Figure 1 shows a landscape of the OpenStack ecosystem where 250 companies have contributed to 817 projects in the 14th release (the most

active release in the development history of OpenStack), resulting in eight contribution models, three collaboration patterns, and four task preferences.

Commercial participation is prevalent and extensive

In the 18 versions of OpenStack from its launch to August 2018, 490 companies have been involved in the

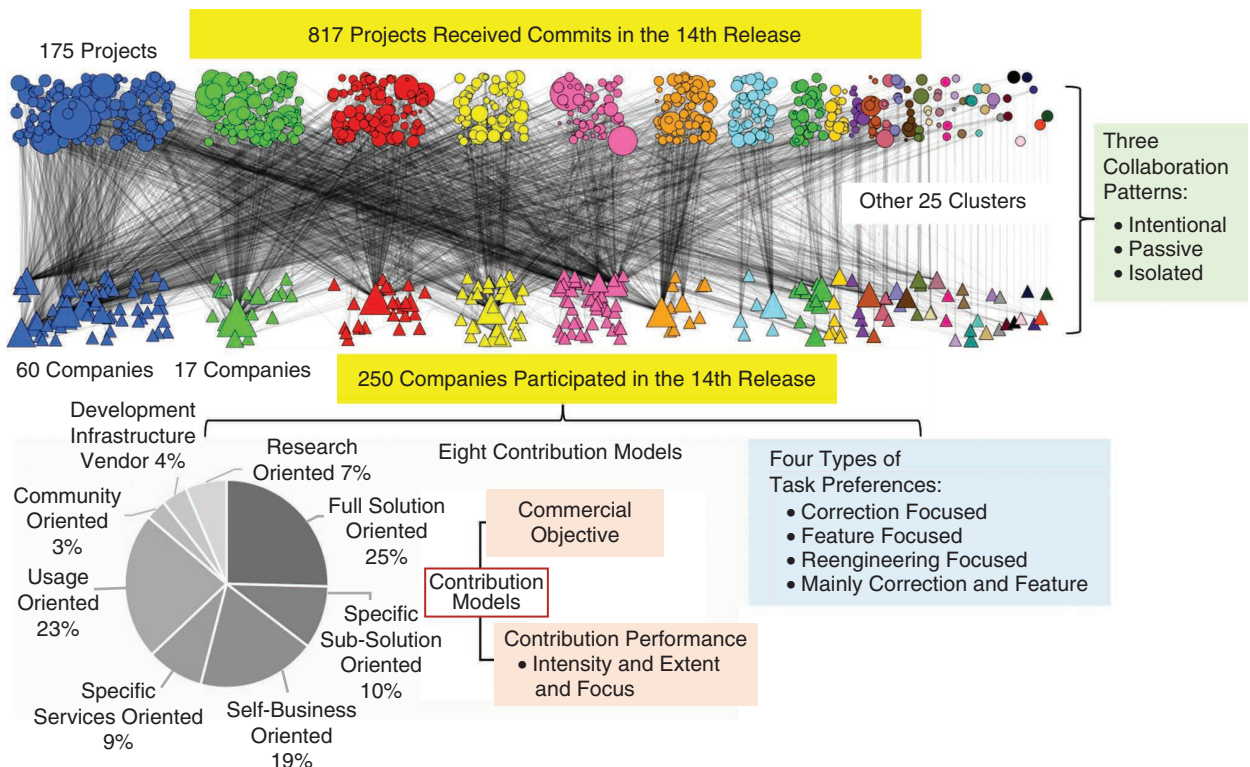


FIGURE 1. The commercial participation landscape in the 14th release of OpenStack.

development. On average, each version has 130 involved companies. Companies made far more contributions than volunteers and have played a critical role in the development of OpenStack, which is evidenced by the amount of developers, commits, and tasks.

The average proportion of developers invested by the companies is

OpenStack, ranging from hardware manufacturers like Dell, semiconductor chip manufacturers like Intel, software providers like Red Hat, to e-commerce platforms like eBay. They join OpenStack because of varying commercial needs, select diverse projects and tasks to contribute to, and form different contribution mod-

through OpenStack, they tend to make intensive and extensive contributions to various OpenStack projects. By contrast, companies select specific projects and make limited contributions when they are motivated by integration aim or are minor contributors (community players or research groups).

While boosting software development, commercial participation also brings challenges and risks to the long-term development of OSS and the sustainability of critical open source ecosystems.

Companies may engage in intentional or passive collaborations or may work in an isolated fashion

Companies are often specialized in one specific domain, and their particular knowledge and expertise have great value in an OSS ecosystem. Collaboration among different companies therefore greatly facilitates efficient development and progress of OSS ecosystems. Most companies in OpenStack collaborate with others, even with their competitors. Some collaborations are actively and intentionally pursued by supplying production deployment, maintaining the same distribution, delegating services, and so on. For instance, Walmart relies on Rackspace for its OpenStack solution, and both companies make extensive contributions to project Ansible founded by Rackspace for deploying OpenStack distributions. However, some company collaborations in OpenStack are passive and incidental. For example, Hewlett-Packard and IBM provide computing services in their OpenStack-based solutions, so both make contributions to Nova, a computing infrastructure project in OpenStack. This collaboration pattern is common in the development of large OSS ecosystems.

approximately 80%, and the proportion of commits made by these developers is 90% in a release of OpenStack.⁹ On the other hand, Figure 2 shows the comparison of different development tasks conducted by companies and volunteers in the 14th release of OpenStack. We can observe that most (91% on average) activities, regardless of their types, are carried out by companies, meeting the Pareto distribution.

els. We discovered eight contribution models in OpenStack⁹ according to the commercial objectives of companies. Each model's participation performance, that is, the intensity, extent, and focus of companies' contributions to OpenStack, varies with its commercial needs.

For example, some companies (full-solution providers) make profits directly by providing full-cloud solutions to users, including private/public/hybrid cloud services, and deployment and maintenance services, among others. As another example, some companies use OpenStack in their production environment or "live symbiotically off an open source ecosystem."¹⁴ When companies obtain profits directly

Various contribution models motivated by commercial needs present different levels of contribution performance

Hundreds of companies with diverse business domains have joined

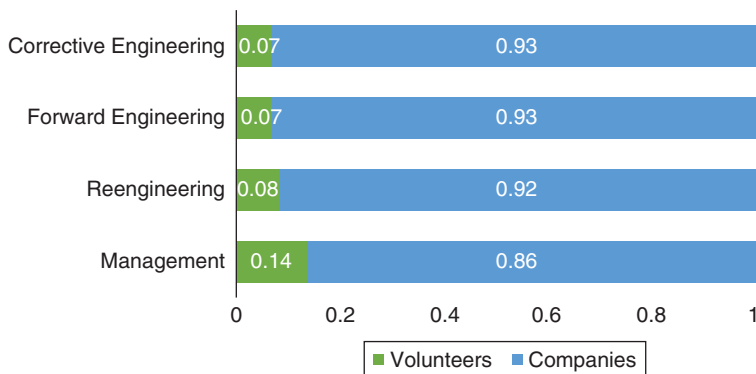


FIGURE 2. The percentage of development activities conducted by companies and volunteers.

during the development of the following six releases. These projects appear to just serve as “an open code archive,” while the true advantages of open source (such as user innovation) are not brought into full play in some corners in a large OSS ecosystem. As a higher degree of collaboration always comes with more efficient development,¹² we see room for improving efficiency in these solo-contributed projects.

Companies have different preferences in development tasks when making contributions

Multiple development tasks such as fixing bugs and implementing new features need to be completed during the preparation of releasing an OSS distribution.¹⁵ A combination of keyword searching and a clustering technique identifies four company task completion patterns, as described in Table 1: focused on correction, feature, or reengineering, or mainly focused on correction and feature. The most common pattern is corrective engineering, in which a company devotes approximately 95% of its commits to fix defects. The reason might be “scratching one’s own itch”;¹⁶ that is, companies encounter a problem of OpenStack, write code to fix it, and send a commit to the community.¹⁷ This is the most common motivation of these casual contributors in OSS, and it may suggest the unsustainability of these companies’ participation in OSS communities. Up to a third of companies’ preferences in the three tasks is close to the general distribution of these tasks in OpenStack, that is, mainly fixing bugs and implementing new features and also doing a few reengineering tasks.

This pattern indicates that the business strategies of these companies toward OpenStack are consistent with the development route of OpenStack or do not affect their choices of conducting development tasks. Some companies focus mainly on adding new features because of the creation of new projects or companies’ specific

objectives toward OpenStack, for example, integrating OpenStack with their own business by contributing plug-ins or drivers.¹² A few companies conduct mainly reengineering activities, such as rename, in some projects. Specifically, 89 out of 222 companies (accounting for 38%) have one

diversity, which is key to the health and vitality of an OSS community.¹⁸

We conducted a survey with developers in OpenStack about the impact of company domination on the development of OSS. More than half (18 out of 34) of the respondents believe that a company’s domination is negatively

To achieve a win-win situation on the sides of both companies and OSS ecosystems, we encourage companies to take into consideration the impact of their participation on OSS projects and delicately balance their own profits with the sustainable evolution of the OSS ecosystems.

preference in development tasks when making contributions to the projects of OpenStack. Companies with a single task preference tend to be short-term contributors, as evidenced by Spearman’s correlation.

The dilemma of commercial participation and the OSS community

Commercial participation can be a two-edged sword for OSS ecosystems. Although the diversity of commercial participation in an OSS project is positively related to the number of volunteers,⁹ the reality is that most OpenStack projects rely heavily on a single company that contributes more than half of commits to these projects. Having one company be the main contributor, that is, domination,⁷ may threaten

related to the number of other companies involved. Some respondents expressed their opinions on the impact of the domination; for example, “high domination results in less involvement from minor players who are concerned that they are free labor to the main sponsor of the project.” The significantly negative associations between a company’s domination degree and the number of other involved companies and volunteers of the projects in OpenStack⁷ have provided statistical evidence for the impact of company domination on the development of OSS. An OSS project that is dominated by a single company for a long time will lose the benefits of collective intelligence and face a high survival risk when the dominated companies withdraw.

TABLE 1. Companies’ preferences in development tasks.

Cluster ID	Task preferences			Count
	Corrective engineering	Forward engineering	Reengineering	
1	0.95	0.03	0.01	1,648
2	0.53	0.38	0.09	1,437
3	0.02	0.97	0.01	792
4	0.01	0.01	0.98	357

Companies are always driven by profit. Once a company achieves its business objectives toward OpenStack or realizes that its objectives can never be satisfied, it may simply withdraw its developers from OpenStack. For example, one developer from Aeode explains why the company withdrew from OpenStack: “Our commercial efforts to make a public cloud were ultimately unsuccessful.” Such withdrawals are not uncommon: more than half of the companies that joined in a certain version of OpenStack withdrew later. The high withdrawal rate of companies can result in losses of code quality¹⁹ (because the historical code contributed

but it may also induce risks for the sustainability of the community in the long term. We call for care and interventions to be taken for the sake of society as a whole.

Companies: A delicate balance in the profits

When joining OSS development or adjusting their existing OSS-participation strategies, companies can follow certain rules, such as the eight contribution models⁹ and three collaboration patterns.¹² These rules start from companies’ objectives toward OSS projects and end with how these organizations select projects, development

To sustain itself, an OSS community should place itself in the whole landscape of the OSS ecosystem, assist the participation of companies and volunteers, and guard the development and maintenance of each project.

by those companies will be left unattended) and productivity (because of the frequent turnover of experienced developers). Moreover, the withdrawal of the dominating companies can also lead to the ultimate failure of an OSS project. For those projects maintained by a single company, the survival risk is particularly severe. For example, the project *quark* was launched by Rackspace in the seventh version to achieve scalable networking service in OpenStack and has become inactive since Rackspace withdrew from it in the 15th version. Quitting halfway not only has a negative impact on OSS communities, but it also wastes the company’s early investment of time and human resources. However, designing a long-term engagement strategy is challenging.

BALANCING COMMERCIAL PARTICIPATION AND OPEN SOURCE SUSTAINABILITY

As shown, commercial participation can make an OSS community flourish,

tasks, and collaborators. A company can design its OSS participation strategies by combining these rules with its business objectives toward the OSS projects. For example, if the objective is to use the OSS product in a production environment, a company should extensively contribute to an OSS ecosystem (to keep up with the development of the projects) and build collaborative relationships with the OSS-based solution providers (to get a timely response). Our measurements of companies’ contribution performance (contribution intensity, extent, and focus), collaborative relationship, and task preferences can be used as monitoring frameworks for these organizations to realize their roles in the development of an OSS ecosystem and learn from other companies to improve the efficiency and quality of their OSS contributions.

However, we must admit that OpenStack is becoming less active: the number of companies contributing to each version decreased from 222 in the 14th

version to 86 in the 24th version. Besides occurring for technical reasons, this decline may also indicate the weakness of commercial participation in OpenStack. For example, a company may dominate the development of one OSS project for maximizing profits or achieving objectives toward the OSS project in a rush. Specifically, the development of many (70%) projects in each version may be contributed mainly by one company. This company domination tends to negatively affect the participation of other companies or volunteers, resulting in two possible cons for the OSS projects: 1) a loss of the benefits of user innovation and 2) a high sustainability risk once the dominating company withdraws from the project. This indicates that the lessons learned from OpenStack cannot be copied directly, and companies should balance their short-term benefits and the long-term sustainability of the OSS ecosystems during the process of making contributions.

Companies can set alarm lines in the monitoring framework to decide when to adjust. Being contributed and maintained entirely by one’s own developers should serve as a warning of not attracting other parties’ contributions. Another possible alarm line could be scarce collaboration because more collaboration with other companies leads to a higher development efficiency.¹² Being focused on a single task could also be a possible alarm because preferring one type of task may lead to a short-term contribution. To achieve a win-win situation on the sides of both companies and OSS ecosystems, we encourage companies to take into consideration the impact of their participation on OSS projects and delicately balance their own profits with the sustainable evolution of the OSS ecosystems.

The OSS community: Governance of commercial participation

To sustain itself, an OSS community should place itself in the whole landscape of the OSS ecosystem, assist the participation of companies and

volunteers, and guard the development and maintenance of each project. Diversity is one of the reasons for the success of OSS,²⁰ and the diversity of companies in OSS projects is proved to have a positive correlation with the participation of volunteers, who can play a mitigation role against companies following tech fashion or hype curves.⁹ OSS communities can reuse our diversity and domination metric of company participation and take action if volunteers become less involved.

The successful onboarding of individual contributors in OSS projects is difficult.²¹ Companies may also face the same challenge: those that conduct mainly a single type of task tend to contribute to no more than one version. OSS communities may identify the companies that prefer to perform a specific development task and find ways to extend their preferred types of tasks. We expect early intervention to increase the stickiness of OSS projects to these companies.

One signal portending an OSS project's failure is that the developers from only a single company maintain it. We noticed that the proportion of companies that become long-term contributors is extremely low. Therefore, an OSS community may need to identify these single-company-maintained projects first, assess their importance, and rescue them (if important) by involving other interested companies. Specifically, an OSS community can reuse our metric of company domination to monitor the health status of its projects. Once the degree of domination is larger than 50% (an empirical value), the OSS community may need to take further steps for assurance.

To harvest the strength of open source, for example, user innovation and low R&D cost, commercial companies have been actively joining OSS projects to make contributions. Driven by their business needs, companies vary on their contribution intensity, extent, and focus, together

establishing a vivid ecosystem as we observed in OpenStack. However, the profit-seeking nature of companies inevitably brings risks to the community. They cannot help controlling the development direction or dominating a project and, therefore, may harm the diversity of the community and affect the inflow of volunteers. They may also withdraw from the dominated project and leave it to die. To balance commercial participation in open source, we propose rules for companies and communities to follow. In particular, they may carefully monitor contributing activities and take actions at the alarm line. What an open source community would be like under intense and differentiated commercial participation will be a controversial topic for a long time, requiring substantial attention in the future. **■**

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of the National Natural Science Foundation of China under grants 61825201 and 62141209.

REFERENCES

1. M. Zhou, Q. Chen, A. Mockus, and F. Wu, "On the scalability of Linux kernel maintainers' work," in *Proc. 2017 11th Joint Meeting Found. Softw. Eng.*, ACM, pp. 27–37, doi: 10.1145/3106237.3106287.
2. N. Harutyunyan, "Managing your open source supply chain—why and how?" *Computer*, vol. 53, no. 6, pp. 77–81, 2020, doi: 10.1109/MC.2020.2983530.
3. D. Harhoff, J. Henkel, and E. Von Hippel, "Profiting from voluntary information spillovers: How users benefit by freely revealing their innovations," *Res. Policy*, vol. 32, no. 10, pp. 1753–1769, 2003, doi: 10.1016/S0048-7333(03)00061-1.
4. J. West and S. Gallagher, "Challenges of open innovation: The paradox of firm investment in open-source software," *R&d Manage.*, vol. 36, no. 3, pp. 319–331, 2006, doi: 10.1111/j.1467-9310.2006.00436.x.

5. J. Feller and B. Fitzgerald, "A framework analysis of the open source software development paradigm," in *Proc. 21st Int. Conf. Inf. Syst.*, Association for Information Systems, 2000, pp. 58–69.
6. M. Zhou, A. Mockus, X. Ma, L. Zhang, and H. Mei, "Inflow and retention in OSS communities with commercial involvement: A case study of three hybrid projects," *ACM Trans. Softw. Eng. Methodol. (TOSEM)*, vol. 25, no. 2, pp. 1–29, 2016, doi: 10.1145/2876443.
7. Y. Zhang, X. Tan, M. Zhou, and Z. Jin, "Companies' domination in floss development – An empirical study of OpenStack," in *Proc. ICSE '18 Companion, 40th Int. Conf. Softw. Eng. Companion*, May 27/ Jun. 3, 2018, pp. 440–441, doi: 10.1145/3183440.3195047.
8. S. Butler et al., "On company contributions to community open source software projects," *IEEE Trans. Softw. Eng.*, vol. 47, no. 7, pp. 1381–1401, Jul. 1, 2021, doi: 10.1109/TSE.2019.2919305.
9. Y. Zhang, M. Zhou, A. Mockus, and Z. Jin, "Companies' participation in OSS development—an empirical study of OpenStack," *IEEE Trans. Softw. Eng.*, vol. 47, no. 10, pp. 2242–2259, Oct. 1, 2021, doi: 10.1109/TSE.2019.2946156.
10. B. Fitzgerald, "The transformation of open source software," *MIS Quart.*, vol. 30, no. 3, pp. 587–598, 2006, doi: 10.2307/25148740.
11. D. C. Mueller and J. Cubbin, *The Dynamics of Company Profits*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
12. Y. Zhang, M. Zhou, K.-J. Stol, J. Wu, and Z. Jin, "How do companies collaborate in open source ecosystems? an empirical study of OpenStack," in *Proc. IEEE/ACM 42nd Int. Conf. Softw. Eng. (ICSE)*, 2020, pp. 1196–1208, doi: 10.1145/3377811.3380376.
13. J. M. Gonzalez-Barahona and G. Robles, "Trends in free, libre, open source software communities: From volunteers to companies," *it-Inf. Technol.*, vol. 55, no. 5, pp. 173–180, 2013, doi: 10.1515/itit.2013.1012.
14. J. Lerner and J. Tirole, "Some simple economics of open source," *J. Ind.*

- Econ.*, vol. 50, no. 2, pp. 197–234, 2002, doi: 10.1111/1467-6451.00174.
15. A. Mockus and L. G. Votta, “Identifying reasons for software changes using historic databases,” in *Proc. Int. Conf. Softw. Maintenance*, 2000, pp. 120–130, doi: 10.1109/ICSM.2000.883028.
 16. E. Raymond, “The cathedral and the bazaar,” *Knowl., Technol. Policy*, vol. 12, no. 3, pp. 23–49, 1999, doi: 10.1007/s12130-999-1026-0.
 17. Z. Li, Y. Yu, M. Zhou, T. Wang, and H. Wang, “Redundancy, context, and preference: An empirical study of duplicate pull requests in OSS projects,” *IEEE Trans. Softw. Eng.*, early access, 2020, doi: 10.1109/TSE.2020.3018726.
 18. A. Bosu and K. Z. Sultana, “Diversity and inclusion in open source software (OSS) projects: Where do we stand?” in *Proc. 2019 ACM/IEEE Int. Symp. Empirical Softw. Eng. Meas. (ESEM)*, pp. 1–11, doi: 10.1109/ESEM.2019.8870179.
 19. A. Mockus, “Organizational volatility and developer productivity,” in *Proc. ICSE Workshop Socio-Technical Congruence*, 2009.
 20. B. Vasilescu, D. Posnett, B. Ray, M. G. van den Brand, A. Serebrennik, P. Devanbu, and V. Filkov, “Gender and tenure diversity in GitHub teams,” in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, 2015, pp. 3789–3798, doi: 10.1145/2702123.2702549.
 21. X. Tan, M. Zhou, and Z. Sun, “A first look at good first issues on GitHub,” in *Proc. 28th ACM Joint Meeting Eur. Softw. Eng. Conf. Symp. Found. Softw. Eng. (ESEC/FSE 2020)*, pp. 398–409, doi: 10.1145/3368089.3409746.

YUXIA ZHANG is an assistant professor at Beijing Institute of Technology, Beijing, 100081, China. Contact her at yuxiazhang@bit.edu.cn.

HAO HE is a Ph.D. candidate at the School of Computer Science, Peking University, Beijing, 100871, China. Contact him at heh@pku.edu.cn.

MINGHUI ZHOU is a professor in the School of Computer Science, Peking University, Beijing, 100871, China. Contact her at zhmh@pku.edu.cn.



IEEE TRANSACTIONS ON BIG DATA

▶ SUBSCRIBE AND SUBMIT

For more information on paper submission, featured articles, calls for papers, and subscription links visit: www.computer.org/tbd

TBD is financially cosponsored by IEEE Computer Society, IEEE Communications Society, IEEE Computational Intelligence Society, IEEE Sensors Council, IEEE Consumer Electronics Society, IEEE Signal Processing Society, IEEE Systems, Man & Cybernetics Society, IEEE Systems Council, and IEEE Vehicular Technology Society

TBD is technically cosponsored by IEEE Control Systems Society, IEEE Photonics Society, IEEE Engineering in Medicine & Biology Society, IEEE Power & Energy Society, and IEEE Biometrics Council

